

Moral Blackmail*

Terrance C. McConnell

I

As moral agents we must deal with the fact that many of our ethical decisions are made under imperfect conditions, and these imperfections may alter our actual obligations. If a person has done something wrong, he may incur a duty of reparation or a contrary-to-duty imperative; if an agent knows that he is weak-willed and will fail to do the morally ideal thing, his actual obligation may be altered so that he is only required to do the best that he can; and if, because of nonculpable ignorance, an agent cannot know that a certain act would be the morally ideal one, he may have no obligation to perform that act.¹ In each of these cases, the imperfection that alters the agent's obligation is his own—the fact that he has done a wrong act, the fact that he is weak willed and will not be able to avoid temptation, and the fact of his own ignorance. It would seem, though, that the imperfections of others might also alter one's obligations. Surely others' moral weakness affects one's own obligations; for example, one has an obligation not to leave confidential material lying on one's desk lest others be overcome by incontinence and look at it. Similarly, it would seem that the ignorance of others may affect one's obligations. At the very least, one might have an obligation to eliminate the ignorance by informing others.

The intuitions are less clear, however, when we ask if the immoral actions of others might alter one's actual obligations. In particular, suppose that one agent, A, threatens to do some evil thing unless a

* An earlier version of this paper was read at the University of Western Ontario and at Wake Forest University. I received many helpful comments from members of those audiences and from others who read this paper. I am especially indebted to Alan Donagan, Barry Hoffmaster, Ralph Kennedy, John King, Arnold Levison, Jan Narveson, John Stevens, and Laurence Thomas for their suggestions.

1. The issue of contrary-to-duty imperatives was introduced by Roderick M. Chisholm in "Contrary-to-Duty Imperatives and Deontic Logic," *Analysis* 24 (1963): 33–36. The question of whether one's own moral weakness can alter one's actual obligations is dealt with extensively by Holly S. Goldman in "Dated Rightness and Moral Imperfection," *Philosophical Review* 85 (1976): 449–87. Whether one's own ignorance can affect one's obligations is discussed by many philosophers. For one example, see H. A. Prichard, "Duty and Ignorance of Fact," *Moral Obligation* (Oxford: Oxford University Press, 1968), pp. 18–39.

Ethics 91 (July 1981): 544–567

© 1981 by The University of Chicago. 0014-1704/81/9104-0007\$01.00

second agent, B, does some less evil thing. Supposing that A has the power to and will carry out the threat unless B complies, does B have an obligation to do evil in order to prevent A's doing a greater evil? Ordinarily, of course, if one is forced to choose among evil alternatives, it is commonplace to say that one ought to do the least of the evils. What seems worrisome about the case in question, however, is that the contemplated action is ordinarily wrong and is presented as a possibly permissible alternative only because of someone else's wrong acts. Can one be morally required to do a normally evil thing simply because, unless one does, someone else will do something worse? Is the notion of "negative responsibility" in this context a morally sensible one?² That we are at least suspicious that obligations can be altered by the immoral actions of others is made manifest by the fact that we are tempted to describe such cases as "moral blackmail."

The question at issue here might be better illustrated if a specific case were presented.

CASE 1: Certain citizens of country K have escaped and come to our country. They fled because of the repressive government of their native land. Believing that they are safe within our borders, they begin condemning the activities of the tyrant of K in order to call public attention to these atrocities. The tyrant cannot tolerate criticism, however, and he demands that our government seize these critics and turn them over to him. He threatens (and we have every reason to believe that he will carry out the threat) that unless this is done, his police will seize the members of our diplomatic staff in his country, of which there are many, and torture them. What should our government do?³

It is reasonable to say that normally it would be wrong to return these few critics to the tyrant, since to do so will ensure that they will be tortured or killed. But in case 1, if we do not return them, a greater evil will be done, namely, more people will be tortured. In situations of this sort we are in a unique position to prevent someone from doing a greater evil, but only by doing a lesser evil ourselves. We have been placed in this unique position by someone else, someone who has directed his threat at us. Let us call situations of this sort cases of moral blackmail.⁴

2. The concept of negative responsibility is discussed critically by Bernard Williams in "A Critique of Utilitarianism," in *Utilitarianism: For and Against*, ed. J. J. C. Smart and Bernard Williams (Cambridge: Cambridge University Press, 1973), pp. 93–118. For a defense of this notion, see John Harris, "The Marxist Conception of Violence," *Philosophy and Public Affairs* 3 (1974): 192–220.

3. Joel Feinberg discusses a case similar to this in *Social Philosophy* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973), p. 87.

4. Some will say that a case of moral blackmail is an example of a genuine moral dilemma, that is, a situation in which an agent ought to do each of two acts both of which he cannot do. In case 1, it might be argued, no matter what the government does it will have done something wrong. In my "Moral Dilemmas and Consistency in Ethics," *Canadian Journal of Philosophy* 8 (1978): 269–87, I have argued that an adequate moral theory must

II

We must ask, then, whether the immoral actions of a moral blackmailer can alter our actual obligations. There are two extreme theories on this question. The first theory states that one is never permitted to do something evil in order to prevent someone else from doing a greater evil. Kant seemed to have endorsed this first view (and hence I shall call it the Kantian theory) when he argued that one is not even permitted to lie to a would-be murderer concerning the whereabouts of his intended victim.⁵ We might also call this the "clean-hands" thesis, since it asserts that one may not dirty one's own moral hands in order to prevent another from doing a greater wrong. Advocates of this theory are committed to the claim that there is a relevant difference between doing evil and allowing evil, and that one is not responsible for allowing evil if it can only be prevented by doing evil. The second theory holds that whenever one must do or allow some evil, one should always choose the lesser evil. If evil will occur no matter what one does, one's obligation is to see to it that the state of affairs involving the least possible evil is brought about. We might call this the "dirty-hands" thesis, since it is claimed that an agent is sometimes obligated to dirty his own moral hands in order to prevent another person from doing a greater evil.⁶ On this view what is ethically important is not the agent's moral purity or integrity, but rather the value of the state of affairs brought about. Thus an agent is just as responsible for forbearing to prevent a greater evil as he is for doing an unnecessary evil. Utilitarians, of course, support this view, but some nonutilitarians do too.

Though each of these extreme views has a number of advocates, neither seems to be plausible. In general, I think that one should be suspicious of normative principles that are as unqualified as these extreme ones are. But in this case we need to examine the particular theories. The Kantian view or clean-hands thesis is suspect because it ignores certain morally relevant features and thus sometimes yields counterintuitive judgments. In some cases, many people do think that one ought to do evil in order to prevent another person from doing

rule out the possibility of genuine dilemmas. I shall assume, then, that whatever we say about cases of this sort, they do not present us with an irresolvable moral conflict.

5. See Immanuel Kant, "On a Supposed Right to Tell Lies from Benevolent Motives," in *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*, ed. T. K. Abbott (London: Longmans, Green & Co., 1909), pp. 361–65. It may be somewhat misleading to call this view the "Kantian" theory since some commentators argue that Kant's ethical system did not commit him to this extreme position. On this point, see H. J. Paton, "An Alleged Right to Lie: A Problem in Kantian Ethics," *Kant-Studien* 45 (1953–54): 190–203.

6. Those who defend or seem to be committed to this view include Kai Nielsen, "Against Moral Conservatism," *Ethics* 82 (1972): 219–31; Michael Walzer, "Political Action: The Problem of Dirty Hands," *Philosophy and Public Affairs* 2 (1973): 160–80; and Bas C. van Fraassen, "Values and the Heart's Command," *Journal of Philosophy* 70 (1973): 5–19.

something worse.⁷ Ironically, Kant's own case seems to be a counterexample to the Kantian thesis.

CASE 2: Person A asks you where B is now. You know where B is, but it is obvious that A intends to kill or do serious harm to B. The only way that you can prevent this injury to B is by lying to A. Are you permitted to tell a lie?

Surely what the agent ought to do in this situation is to lie to A, even though lying is normally impermissible. In fact, this case is often cited to show that there are exceptions to moral rules or that the same action-type can be right in one situation but wrong in another.⁸ Another case that casts doubt on the Kantian view is taken from Shakespeare's *Measure for Measure*.

CASE 3: Angelo, the head of the city's government, unjustly condemns to death one of his subjects, Claudio, for the crime of lechery. Isabella, Claudio's sister, goes to Angelo and pleads for her brother's life. She is a devoutly religious person. Angelo agrees to free Claudio, but only on the condition that Isabella have sexual intercourse with him. So apparently she can save her brother's life only by allowing herself to be sexually exploited.

I have not described this case entirely accurately. In the play Isabella does find a way to save Claudio without yielding to Angelo. However, it is still important to consider this case because when first confronted with the decision, Isabella has reason to believe this is the only way to save her brother; in addition, the case I describe is just as possible as the one that unfolds in the play. Whether Isabella is morally required to succumb to Angelo's demands is not obvious. Allowing oneself to be treated in such a way does, after all, involve a great sacrifice. It does seem reasonable to say, however, that it would not be wrong for Isabella to take these steps in order to save Claudio. I say this, not because of a liberal attitude toward sexual morality, but because of what will be lost if she fails to take these steps, namely, a man's life.⁹ Since the first theory entails that such an act would be wrong (because Isabella would be "giving in" to evil), one has reason to believe that the theory is inadequate.

Let us consider one other case that raises problems for the clean-hands theory, a case that is not hypothetical.

CASE 4: The Hanafi Muslims seize control of a public building in Washington and hold 130 people as hostages. They threaten to kill

7. Here and elsewhere in the paper I do not mean to suggest that there is a common set of intuitions concerning the cases I discuss. Certainly there is no such consensus. In justifying the particular judgments that I make, I shall call attention to considerations that are morally relevant but which are ignored by the extreme theories.

8. See Paul W. Taylor, *Principles of Ethics* (Belmont, Calif.: Dickenson Publishing Co., 1975), p. 2.

9. In short, I (and I believe many other people) put a higher value on preventing an unjust killing than I put on resisting sexual exploitation. I realize that, as a sixteenth-century Catholic, Isabella's views may have been somewhat different.

all of the hostages (and we have every reason to believe that they will) unless their demands are met. They demand that the movie *Mohammad, Messenger of God* be taken off the market because, they say, it is blasphemous. Should we agree to their demand?

Suppose that this were the only demand that the terrorists made. If it were, many would agree that the movie ought to be temporarily removed from the market in order to save the lives of 130 people. This is so even though it is normally wrong to restrict people's freedom or to engage in censorship. Again, the Kantian view denies that we ought to give in to the terrorists' demands. Certainly, according to that theory, one may not harm innocent persons in order to prevent another from doing a greater evil. This, I think, casts serious doubt on the view.

This may lead one to believe that the second theory, the dirty-hands thesis, is correct. This theory does yield judgments that a number of people are inclined to endorse in cases 2, 3, and 4. However, the theory yields less palatable judgments in some other cases. Consider a more detailed and historically accurate description of case 4 (which I shall call the extended version of case 4). The Hanafi Muslims make an additional demand—that the five Black Muslims convicted for the massacre of a family at the Hanafi headquarters be turned over to them so that they can “settle the score” or “get their justice.” If the Black Muslims are given to the Hanafis, we know (let us assume) that they will be killed; but if we fail to do this, we know that many of the 130 hostages will be killed. The second theory, then, would require us to hand over the persons; in this case, the loss of five lives is the lesser evil.¹⁰ Yet it can be argued that giving in to the demand is wrong. Not only does it make a mockery of the ideal of rule by law, but doing something which one knows will lead to the torture and death of others under such circumstances is surely impermissible. In this case, numbers do not matter. It seems reasonable to say, then, that this demand ought to be rejected. The implications of the second theory are even more worrisome in the following case.

CASE 5: Person A holds a gun at your head. He threatens to kill you and your two children (and you have every reason to believe that he will) unless you kill person B, an innocent party against whom A holds a grudge. May you kill B in order to save yourself and your children?

As the case is described, no matter what the agent does, evil will occur. If the agent does nothing, three people will be killed. But if the agent complies with A's command, only one person will be killed. Since the lesser evil is the death of one person, the second theory requires that the

10. One qualification must be made here. We must assume that this act will not lead to more evil—for example, the loss of even more lives—because of the precedent that it sets. As D. H. Hodgson points out in *Consequences of Utilitarianism* (Oxford: Clarendon Press, 1967), pp. 77–78, there is a real danger that we shall encourage acts of this sort by complying with the demands. We can imagine, of course, that the negotiations in case 4 take place in private and that the settlement is not known to the public.

agent dirty his own hands in order to prevent A from doing a greater evil. This, however, seems morally unacceptable. Since in this case the agent must kill B in order to save his own life, some might describe the act as one of self-defense. But even if this were granted, there are moral limits on what one may do to defend oneself. The mere fact that A's continued existence poses a threat to the life of B does not justify B's taking the life of A. Person A may be an innocent party, and it is hard to see how killing an innocent person can correctly be described as a case of justified self-defense.¹¹ Once again we have reason to doubt the adequacy of the second theory.

The problem of how cases of moral blackmail may affect one's actual obligations is now clear. There are some situations where it can plausibly be argued that one ought to do something evil in order to prevent the moral blackmailer from doing a greater evil. There are other situations, however, where no matter how great the evil is that one can prevent, what one must do in order to prevent it is impermissible. The Kantian theory does not allow for cases of the first kind, and the dirty-hands thesis rules out the possibility of situations of the latter sort. Thus each of the extreme views seems to be incorrect. Aristotle, in the *Nicomachean Ethics* (3, 1, 1110a25–30), seems to have recognized the inadequacy of the extreme views. There he suggests that we rightly excuse a person when he is forced by another to do something evil in order to save his own life or the lives of his children. But, he says, some acts are so evil that no matter what the other person threatens to do, one is not permitted to do what he demands. Aristotle suggests that killing one's own mother is an example of such an impermissible act. The challenge, of course, is to formulate a principle to handle cases of this sort. Relying on intuitions will not always do, since there are a number of cases about which we are uncertain. What is needed is a compromise theory, a principle which is free from the defects of the extreme theories and which enables us to resolve the difficult cases. If no such principle exists, some of us will be in the awkward position of being forced to abandon some moral judgments that seem quite defensible, no matter which of the extreme theories we opt for. To the difficult task of constructing such a principle we now turn.

III

There does seem to be something initially appealing about the Kantian view. To all but the act utilitarian, the dirty-hands thesis must seem suspect. But what is it about the extreme version of the dirty-hands thesis that seems implausible? In the extended version of case 4, the theory

11. For a defense of this claim, see Baruch A. Brody, "Abortion and the Sanctity of Human Life," *American Philosophical Quarterly* 10 (1973): 133–39, see esp. p. 134. See also Judith Jarvis Thomson, "A Defense of Abortion," *Philosophy and Public Affairs* 1 (1971): 47–66, see esp. p. 53; and Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 62–63.

requires that we relinquish the five Black Muslims to the Hanafis, even though we know that this will result in their being tortured and killed. This is wrong, at least in part, because the Black Muslims are being treated as a means only, as the Kantian might put it. The situation is similar in case 5. To kill person B would prevent a greater evil—namely, the deaths of three people—but this would be a blatant case of using B as a means only. In each of these cases, if the agent were to do what was necessary to prevent the greater evil, he himself would be violating the rights of some innocent party (whose rights would not be violated if he did nothing).

Yet, as we saw, there are cases that cast doubt on the clean-hands thesis as well. In case 2 it is surely permissible for the agent to lie to A in order to prevent him from killing B. Here A is the “victim” of the agent’s act, but of course he is not an innocent party. In this situation, we do not owe A the truth. A’s rights have not been violated; either he has forfeited his rights or they have been overridden.¹² In case 3 Isabella may comply with Angelo’s demands in order to save Cluadio. Since she is the agent who is allowing evil to occur in order to prevent a greater evil and she is the victim too, she is surely permitted to make such a sacrifice if she wishes. She is not violating one of her own rights (if such talk even makes sense), though of course Angelo is doing something wrong. Case 4 may seem more troublesome. In cases 2 and 3 the agent who does or allows the evil is not violating some innocent party’s rights. But in case 4 the rights of some innocent party are being tampered with, namely, the rights of those who own and are showing the movie. We are harming some innocent party in order to prevent other parties from suffering more. Why is such an act permissible in this case but not in the others? What makes case 4 different from the extended version of that case or from case 5? Though in each of these cases one can prevent a greater evil only by violating the rights of some innocent person, the rights to be violated in the latter two cases seem to be qualitatively more important than the right at issue in case 4. The rights in the latter two cases are ones to which we attach a much greater strength. And with respect to these examples, why we attach these differing strengths seems clear enough. In case 4 we can easily compensate the injured parties for their losses, financial and otherwise. This is not the case, though, when what the victim loses is his life. When the right to life is at stake, compensating the victim for his loss is not an option. This seems to be an important and morally relevant difference.

12. Some who hold that human rights are absolute think that there is a problem with saying that rights can be forfeited or overridden. As an alternative, one might adopt a view suggested by A. I. Meldon in his *Rights and Persons* (Berkeley: University of California Press, 1977), pp. 61–62. Meldon argues that one’s right to truthfulness presupposes that the interests one will pursue are not morally objectionable; that is, one has a right to the truth only if one will not use this information to further one’s pursuit of immoral projects. According to Meldon’s view, the would-be murderer never had the right to be told the truth.

Having seen, at least in part, why the extreme versions of the Kantian theory and the dirty-hands thesis are inadequate, we are in a position to formulate a principle which remedies these inadequacies. Let us consider the following principle (*P*) as an alternative to these views.

An agent in a situation of moral blackmail may do something evil in order to prevent the blackmailer from doing a greater evil just in case either (1) the agent's action does not irreparably violate the rights of some nonconsenting, innocent person, and if an innocent person's rights are violated he is to be compensated, or (2) if a nonconsenting, innocent person's rights are irreparably violated, then the same (or equally strong) rights of that same person would have been violated even if the agent did not comply with the blackmailer's demand.¹³

An explanation of principle (*P*) is in order here. First, (*P*) is a strong principle in that it sets out both necessary and sufficient conditions for when it is permissible to comply with the demands of a moral blackmailer. Next, it is important to note that principle (*P*) allows an agent to violate an innocent person's rights on the condition that it is necessary to prevent the moral blackmailer from doing a greater evil. Some might claim that it is always permissible to violate another person's rights provided that the violator fully compensates the victim. However, it has been argued—persuasively, I think—that this view must be rejected.¹⁴ Principle (*P*) does not give one such a wholesale license to do evil. Third, the qualification “innocent person” is included to allow that evil, even irreparable evil, may be done to a person in order to prevent that same person (or one conspiring with that person) from doing evil to someone else. It is to this intuition that the Kantian view runs counter, for it even forbids one to lie to a would-be murderer. Fourth, the qualification “nonconsenting” is included because if a person chooses to sacrifice himself in order to prevent the blackmailer from doing a greater evil, it is reasonable to say that he may do so, even if the harm he suffers is irreparable. Such an act of martyrdom is not only not impermissible, it seems to be above and beyond the call of duty, and any plausible moral principle must allow for such acts.

The fifth point of explanation concerns the requirement of compensation. Principle (*P*) permits an agent to violate the rights of a

13. In *The Theory of Morality* (Chicago: University of Chicago Press, 1977), Alan Donagan mentions a thesis somewhat similar to clause 1 of principle (*P*). He says, “As a rough generalization, you may, under coercion, be an involuntary accomplice (in the colloquial sense of ‘involuntary’) to a wrong, provided that your action does not irreparably violate what is due to some human being as such” (p. 173). Donagan, however, neither explains nor defends this thesis. Moreover, it is not clear whether he is providing necessary and sufficient conditions for when such acts are permissible, or merely sufficient conditions (depending on how one construes “provided that”). Finally, things that he says elsewhere in the book suggest that he is sympathetic with what I call the Kantian theory. See pp. 146–47, pp. 207–8, and his discussion of the Pauline principle, esp. pp. 149–57.

14. Nozick, chap. 4, esp. pp. 65–71.

nonconsenting, innocent person if that person can be compensated for that right-violation. This clause is necessary in order to avoid the objectionable consequence of the dirty-hands thesis, the consequence that an innocent person may be sacrificed or treated as a means only simply to avoid a greater evil. It may be that principle (*P*) still allows some innocent people to be treated as a means only, but in requiring that amends be made the force of this objection is considerably diminished. Finally, the notion of an irreparable violation of a person's rights must be explained. There are two different senses in which a violation of rights might be irreparable, what I shall call the strong and weak senses. A violation of rights is irreparable in the strong sense if the victim cannot be compensated in any way for his loss. To cite an obvious example, if a person's right to life is violated, no compensation for the victim is possible. If irreversible psychosurgery is performed on a normal person (against his will) transforming him into a raving madman, this violation of rights would be irreparable in the strong sense. A violation of rights is irreparable in the weak sense if the victim can be compensated to some degree, but cannot be adequately compensated. For example, a person may possess a painting which he regards as priceless, perhaps because it was painted by Rembrandt, the person's favorite painter. Similarly, an object may be irreplaceable to a person because of its history, for example, an heirloom. The owner has a right to these objects, and if they are taken from him adequate compensation cannot be made. However, some compensation can be offered. Money will not buy the original Rembrandt that was lost, but it will buy something of similar value. It should be noted that the distinction between the strong and weak senses of irreparability is not based on the actual preferences of the victim; that is, it is not the victim who determines whether any compensation at all is possible. One reason for this is that in some cases we may have to decide whether to comply with the blackmailer's demands without knowing what those preferences are. Another reason is that we may regard those preferences as irrational. For example, we would regard the movie owners (in case 4) as quite unreasonable if they claimed they could not be compensated in any way for the temporary seizure of their film. If a person is alive and well-off enough to enjoy the things he normally does, then it seems reasonable to say that the person can in some way be compensated. Clearly, if a person is killed, permanently turned into a madman, or has all of his limbs severed, he cannot enjoy many of the things that he normally would.

One of the intuitions underlying principle (*P*) is that some rights are so strong that they ought to be protected even if the only means for doing so involves violating someone else's weaker rights. Since we do regard rights the violation of which is irreparable in the strong sense as qualitatively more important than those whose violation is irreparable only in the weak sense, in (*P*) the term "irreparable" is being used in the

strong sense.¹⁵ So principle (*P*) does not protect an innocent person's possessions which he deems priceless. The justification for this is not that such property rights are unimportant; clearly, they are important. But these rights are not as strong as, say, the right to life of other innocent persons. Thus if we must choose between complying with the moral blackmailer's demands and giving him some innocent person's original Rembrandt or allowing him to kill some innocent people, we may take the former course of action if we compensate, to the degree possible, the owner of the painting. (Clause [2] of principle [*P*] will be explained in Section IV.)

As the preceding remarks suggest, one of the strengths of principle (*P*) is that it is free of the defects that plague the extreme theories; it takes account of morally relevant factors ignored by those views. The principle implies that in case 2 it is permissible to lie to person A because he is not an innocent victim. Since he intends to kill B, his right to be told the truth has been overridden or forfeited. Concerning case 3, (*P*) does not forbid Isabella to allow herself to be used as a sexual object in order to free Claudio (assuming that she cannot free him in any other way). Though she is innocent and ought not to be harmed, she may, if she wishes, choose to suffer in order to save her brother. And if we apply the principle to case 4, we may stop the showing of *Mohammad, Messenger of God* in order to save the lives of the 130 hostages. We are required, however, to make amends to those who own and profit from the showing of the movie. In this case the compensation would presumably be monetary. When we consider the extended version of case 4, however, (*P*) forbids us to comply with the demands of the terrorists. Since turning over the prisoners to the Hanafi Muslims will have as a direct consequence their being tortured and killed (we shall assume), to comply would irreparably violate their rights. More to the point, this violation would be ensured by our act, and it is this that (*P*) forbids. For similar reasons, it would be wrong to do the evil demanded of the agent in case 5. If we kill person B, we will have irreparably violated one of his rights. So such an act is forbidden to us, even though someone else will irreparably violate the rights of three people. Thus principle (*P*) takes into account certain moral distinctions that seem quite important, distinctions that the two extreme theories neglect.

There are, however, other reasons for accepting (*P*). This principle is, for the most part, consistent with the natural rights tradition. While no complete analysis of natural rights can be given here, we can say that

15. I have said that a violation of rights is irreparable in the strong sense if the victim cannot be compensated in any way for his loss. A brief explanation is needed here. It must be the case that he cannot be compensated because of the very nature of the right violated and not because of some contingent fact. For example, if I steal a dollar from Jones and he dies in the next hour, I cannot in any way compensate him for the loss of the dollar. But this is the result of a fortuitous circumstance, and not because of the nature of the act.

minimally rights are entitlements, or that rights are claims one has against others. Rights put legitimate restraints on the actions of others; they put a strict limit on what others may do to us. Moreover, the possession of natural rights in no way depends on the political institutions or the social order of which one is a part. When a person possesses rights, this imposes obligations on others. One of the fundamental features of rights is that they protect the individual against utility-maximizing claims.¹⁶ There are certain things that may not be done to the individual, even in the name of maximizing the good. The prohibition against knowingly punishing an innocent person in order to prevent some great evil, such as a riot, is a classic example. If a person were always permitted to do evil in order to prevent someone else from doing a greater evil, the possession of rights would be of little value, or at least their value would be considerably less. Part of the appeal of principle (*P*), then, is that it does, at least in part, protect one's rights against utility-maximizing claims.

There is an objection to this, however. It might be claimed that principle (*P*) ignores the rights of the parties who will be harmed by the immoral actions of others. And since they are the greater number, or at least the evil done to them is greater, this objection seems to have some point. One defender of the dirty-hands thesis states this objection forcefully: "But cannot the compliment [of inhumanity] be returned by speaking of the even greater inhumanity, conjoined with evasiveness, of those who will allow even more death and far greater misery and then excuse themselves on the ground that they did not intend the death and misery but merely forebore to prevent it?"¹⁷ The response to this objection should make the intuition behind principle (*P*) clearer. The rights of those being harmed are not ignored. The advocate of principle (*P*) would of course acknowledge that the rights of the victims have been violated. However, the point is that the rights of other innocent people put a strict limit on what may be done to prevent others from violating someone's rights. Clearly a society should protect its citizens against the crimes in question, such as murder and torture. But just as surely it may not ignore the rights of innocent individuals in trying to fulfill this obligation. This same point can be made another way. Suppose that the moral blackmailer will torture A and B unless the agent in question tortures C. According to (*P*), it would be wrong to comply with these demands. Notice that if the agent were to comply with the demands, C would have a legitimate grievance against that agent. That agent would be the one who irreparably violated C's rights. If the agent does not give in to the demands, however, A and B do not have a complaint against him, or at least their complaint is not of the same magnitude as C's would be. This might suggest (but does not imply) that negative rights (e.g., the

16. This is argued by Alan H. Goldman, "Rights, Utilities, and Contracts," *Canadian Journal of Philosophy* 3, suppl. (1977): 121–35.

17. Nielsen, pp. 229–30.

right not to be killed) take precedence over positive rights (e.g., the right to be protected against harm) in cases of conflict.¹⁸ There is, however, no need to endorse this general thesis here.

However, principle (*P*) does not go as far as some defenders of the natural rights tradition. Most notably, it does not go as far in defending natural rights as libertarians do. Some would say that a person's rights may never be violated, even if compensation is paid, unless the person consents to the act (and then it is no longer a violation of his rights). Robert Nozick considers the hypothesis that all violations of rights (what he calls "border crossings") should be prohibited, but rejects this because some boundary crossings with compensation pay their own way; that is, their benefits are very great. But Nozick still places strict limits on when such boundary crossings may occur. "Any border-crossing act which permissibly may be done provided compensation is paid afterwards will be one to which prior consent is impossible or very costly to negotiate. . . ."¹⁹ On Nozick's view, then, even if one sufficiently compensates a person for a border-crossing act, that act will be wrong unless the victim's prior consent was impossible or very costly to obtain. One may note how this contrasts with principle (*P*) by examining case 4. This is the case in which the Hanafi Muslims will kill 130 people unless a movie to which they have strong religious objections is removed from the market. Suppose that those who own and profit from the showing of the movie protest. It is, after all, their property. Would we not be violating their rights to seize the movie, even if the confiscation is temporary and the owners are adequately compensated (for their inconvenience as well as loss of profits)? On Nozick's view, taking the movie off the market would be an impermissible boundary crossing. Clearly, obtaining prior consent was possible (i.e., we could easily contact and talk to the owners) and not very costly. So if the owners refuse to stop showing the movie in order to save the 130 hostages, we may not force them to do so. Principle (*P*), however, allows us to force the owners to stop showing the movie (at least temporarily) provided that we adequately compensate them. In this case I believe that principle (*P*) better matches the considered moral judgments of most people than does Nozick's view. So though (*P*) does not go as far in protecting our rights as some would, this is not a weakness but rather a strength of the view.

There is an obvious problem that must be mentioned at this point. It is suggested by my acceptance of principle (*P*) that some acts irreparably violate another person's rights (in the strong sense), while other border crossings are such that the victim can be compensated, sometimes adequately and sometimes only to some degree. But this raises an obvious question. How can one tell when a violation of a person's rights

18. This thesis is defended by Philippa Foot in "The Problem of Abortion and the Doctrine of Double Effect," in *Moral Problems*, ed. James Rachels, 2d ed. (New York: Harper & Row, 1975).

19. Nozick, p. 72.

is compensable and when it is not compensable? What is the criterion for determining whether amends can be made? Furthermore, how does one tell what sort of compensation is appropriate if one admits that adequate compensation cannot be made? No answer to these difficult and important questions will be provided here. The examples given, though, do show that there are clear instances of all three types of cases: ones in which the person whose rights are violated can be fully compensated, ones in which he can be compensated only to some degree, and ones in which the victim cannot be compensated at all. One might infer from the examples given that property rights are compensable and that personal rights are not. Such an account, however, is surely too simple.²⁰ It may be that a complete theory of rights is needed before one has a principled way of determining when compensation can be made and when it cannot. The task of providing such a theory, however, is not one that I shall pursue in this paper.

IV

The last clause of principle (*P*) must now be explained. This clause states that one is permitted to violate irreparably the rights of some nonconsenting, innocent person in order to prevent a moral blackmailer from doing an even greater evil just in case the same (or an equally strong) right of that same person would be violated even if the agent failed to comply with the blackmailer's demand.²¹ The reason for this clause is simple enough. If the rights of some person *A* will be irreparably violated even if one does not comply with the demand, then if one's being the agent of the harm done to *A* will prevent a much greater evil, it seems reasonable that that is what one ought to do. Let us consider a case to illustrate this.

CASE 6: Jim, having gotten lost while on a botanical expedition, finds himself in the central square of a South American town. Twenty Indians are about to be killed in order to deter other Indians from protesting against the government. Since Jim is an honored guest, the captain of the soldiers offers him an alternative: if Jim will kill one of the Indians himself, the others will be allowed to go free; if Jim refuses, all twenty will be killed. Short of killing the one, there is nothing that Jim can do to stop the captain. The

20. It might also be tempting to say that in cases of conflict personal rights always override property rights. This view, however, is also much too simple, as is shown in Feinberg, pp. 76–77 (see n. 3 above).

21. Some claim that whenever an agent does an act that is normally evil he does some damage to his moral character. Such an act makes one more callous, it is argued. This is especially so if the act involves irreparably harming some innocent person. Though I think that these claims are considerably exaggerated, I need not take a stand on the issue here. We may simply stipulate that when we consider whether a certain act, *x*, really does prevent a greater evil, our determination of this will include the damage, if any, that the character of the agent of *x* suffers.

twenty and the other villagers understand the situation, and they obviously want Jim to accept. What should he do?²²

If Jim does nothing, all twenty of the Indians will be killed by the captain; however, if Jim kills one of the Indians, the others will be freed unharmed. His only other alternative, trying to seize the weapon and save all of the Indians, will result in his own death as well as that of the twenty. Hence principle (*P*) entails that Jim may kill the one even though this irreparably violates his rights.

In a case such as this, one is tempted to say that there is some sense in which Jim is not really violating the victim's rights. He did not create the situation. Nevertheless since killing is such a serious act, one may wonder how it could ever be permissible. This can best be answered if we look at the case through the eyes of the twenty. In particular, we must ask what sort of complaint the twenty Indians have and against whom. Clearly, they have a legitimate complaint against the captain. It is he who has put their lives in danger without just cause. Because of the actions of the captain, no matter what Jim does, some evil consequences will ensue. But the Indians have no complaint against Jim (assuming, as we do, that there is nothing that he can do to save all twenty). After the captain has made his declaration, the best state of affairs possible is the death of one of the Indians. Since the death of one Indian is inevitable and since if Jim does nothing that same Indian plus the other nineteen will die, Jim's agreeing to execute the one seems to be the right act to do. In fact, each of the twenty desperately wants Jim to accept this "privilege" that honored guests are accorded. In some sense, then, the one killed has consented to Jim's killing him. So even the one Jim chooses to kill (or relatives of that one) will have no complaint against Jim, at least if we assume that he uses some random method to select his victim. The victim has lost nothing that he would not have lost had Jim not complied with the blackmailer's demand. Normally, of course, that the victim consented is neither a justification nor an excuse for murder. But in extreme circumstances such as these, I think that it is reasonable to regard the victim's consent as justificatory grounds for the act of killing. Notice, for example, that some who would ordinarily reject this reasoning approve of voluntary euthanasia when the person to be killed will die soon anyway. Thus, that the victim's same rights will be violated even if Jim does not comply with the demands of the blackmailer seems to provide grounds for altering our judgment significantly.

It is important to note that the last clause of principle (*P*) does not entail that an agent may kill any person who is going to die soon anyway to prevent the blackmailer from doing a greater evil. The nature of the cause of the person's death is morally relevant. Baruch Brody, in discussing the question of when it is permissible to kill an innocent

22. This is a slightly modified version of a case presented by Williams, pp. 98–99 (see n. 2 above).

person, worries about the following type of case: a situation in which we can prevent five people from being blown up by a bomb only by blowing up the bomb's triggering mechanism, thereby killing one innocent person who is not one of the five but who will die soon anyway from cancer. Brody thinks that we have conflicting intuitions about such a case. On the one hand, it seems that killing the person in order to save the five lives is permissible since he will die soon anyway. On the other hand, the nature of the cause of his death seems to be relevant, and this would suggest that killing this innocent person would be wrong. With some hesitation, Brody concludes that what is more important is that the agent will die soon anyway, and therefore killing him to prevent a greater evil is permissible.²³ Brody's account is a general one intended to deal with all cases of permissible killing. I am only concerned with situations of moral blackmail, and in these cases principle (*P*) yields different results than Brody's view. According to principle (*P*) it would be wrong to kill a person or otherwise irreparably violate his rights unless the same rights of that person would be violated even if the agent did not comply with the blackmailer's demands. If the moral blackmailer demanded that we kill some innocent person whose rights would not otherwise be violated but who was going to die soon anyway from natural causes, principle (*P*) would forbid such an action. Dying from natural causes does not violate one's rights.

It is not easy, I think, to choose between principle (*P*) and Brody's view. Let us examine the implications of the views, however, by considering the following.

CASE 7: A small village is surrounded by a group of brigands who demand that the villagers kill Joe, an innocent member of the village whom the brigands dislike. If the villagers fail to comply with this demand, the brigands threaten to (and the villagers have every reason to believe that they will) destroy the village and everyone in it, including Joe. The village is cut off from outside help and giving in to their demands is the only way to save the village.²⁴

Though this is a difficult case, if these really are the only alternatives open to the villagers, I believe that many do think that they are permitted to kill Joe. Brody agrees. But consider a modified version of case 7. Suppose again that the brigands demand that the villagers kill Joe. If they do not comply with the demand, the brigands will kill the five village leaders (and this does not include Joe). And, as it happens, Joe has cancer and will die in six months anyway. In this situation it can be argued that the villagers ought not to comply with the demands of the brigands. Even though giving in to their demands would prevent a

23. Brody, pp. 138–39. This particular case, of course, is not one of moral blackmail.

24. *Ibid.*, p. 136.

greater evil, it is reasonable to say that it would be wrong to do so.²⁵ Joe has a right to live out the remaining few weeks of his life. Of course, we think that Joe would be a splendid person if he would volunteer to have his life ended prematurely in order to save the five leaders. However, this is not something that we may force him to do. If the line of argument presented here is plausible, it lends support to my claim that what is morally important (at least in situations of moral blackmail) is whether the same (or equally strong) rights of the same person will be violated anyway, and not whether the person will die soon.

One serious question remains concerning the last clause of principle (*P*). How can an agent ever know that no matter what he does, the rights of some innocent person will be irreparably violated? How can Jim know, for example, that were he not to comply with the captain's demand the person he executes would have been killed anyway? Some might say that since Jim can never know that he will fail, he should at least try to seize the captain's weapon in order to save all twenty, and he should do this even if it happens to turn out that all twenty-one of them are killed. When the other alternatives open to one are as evil as they are in this case, one should always try to do the ideal act. There are two different ways to respond to this objection. One response is to point out that the last clause of principle (*P*) tells an agent what he ought to do if in fact the same rights of that same person will be violated anyway. The problem of the agent's knowing the relevant facts is independent of the adequacy of principle (*P*). There is, however, another response that one can make. Principle (*P*) can be altered so that it tells an agent what he ought to do given the evidence that is available to him. Thus if an agent had reason to believe that he could save all twenty of the Indians, though in fact he could not, the revised version of (*P*) would allow him to try to save all twenty. Even if this is granted, though, there can still be cases in which the evidence available to the agent indicates that saving all twenty is not an option open to him. In this case, he would have to make the same difficult choice as before, and the revised version of (*P*) would say that he is permitted to kill the one. That agents must sometimes act without knowledge of all the relevant facts is itself a case where moral decisions must be made under imperfect conditions. Since it is desirable not to beg any questions about what an agent's obligations are under such conditions, I shall simply point out that principle (*P*) may be stated in two ways: either an agent's obligations depend on what in fact will occur, or his obligations depend on what he has good reason to believe will occur.

25. This clause of principle (*P*) has the controversial consequence that Joe may not be killed even if all of the other villagers (but not Joe) will be killed by the blackmailer. I accept this consequence, and simply acknowledge that further discussion of the extreme "end of the world" type of situation is needed.

V

Principle (*P*) tells us when it is permissible to comply with the demands of a moral blackmailer. More specifically, (*P*) states under what conditions it is permissible to do evil in order to prevent the blackmailer from doing a greater evil. One must ask, though, whether the fact that the actions of a moral blackmailer are involved presents a special problem. One can imagine a number of situations involving natural disasters such as disease, famine, and drought where an agent can prevent a greater evil only by doing something that is normally wrong. Consider the following (artificial) example.

CASE 8: A huge rock, jarred loose by a tremor, is rolling down the hill. If it continues on its present course, it will kill twenty people. Fortunately the course of the rock can be diverted; unfortunately the only way it can be diverted is such that two, and only two, different people will be killed by it. What should some observer do if he is the only person who has the power to divert the course of the rock?

In this case one can prevent the deaths of twenty people only by doing something that will cause the deaths of two other innocent people. Are cases such as this one any different in principle from those that involve immoral actions on the part of others? Bernard Williams, with his emphasis on the distinction and possible conflict between one's own projects and the projects of others, suggests (but does not develop the point) that these two types of cases are importantly different.²⁶ Critics of Williams, however, have thought that this is absurd. "Other people's projects must, after all, be seen simply as a class of need-creating circumstances and must, for the purposes of moral decision-making . . . , be treated simply as features of a brutal world, like famine and disease."²⁷ There seems to be something appealing about the position of the critic. If a person is about to be harmed and we can prevent it only by harming someone else, what moral difference does it make whether the source of the harm is someone else's wrong actions or nature?

Yet in spite of the apparent plausibility of the critic's claim, it can be argued that there is a moral difference between situations of moral blackmail and those where the source of evil is nature. The following two examples call attention to this moral difference. Consider case 4 again. Surely in that situation it is permissible to seize the movie from its owners in order to save the lives of the hostages (provided that we compensate the owners). But consider an analogous case. Suppose that we could prevent considerable suffering—perhaps even save some lives—by seizing the profits of one who owns a movie and using them to buy food for some suffering from famine. In this case the source of evil is a natural

26. Williams, pp. 109–18.

27. John Harris, "Williams on Negative Responsibility and Integrity," *Philosophical Quarterly* 24 (1974): 265–73, see esp. p. 273.

disaster, yet it seems wrong to alleviate the suffering in this way. Thus there seems to be a moral difference in the two types of cases. It will be helpful if we can say why there is such a difference. On what are these different judgments based? In case 4 we are permitted to seize the movie to protect the stronger rights of the innocent hostages. If we do not do this, they will be killed. Not only do they have a right not to be harmed, but they also have a right to be protected against such border crossings. In the other case, however, what is at issue is the right to be aided or the claim to beneficence. Some (e.g., libertarians) would deny that there is a general right to beneficence. But even if there is such a right, it is reasonable to maintain that this right is not as strong as the right not to be harmed. The difference in strength of these two rights explains, at least in part, why there is a moral difference between these two cases. When the source of evil is nature, the potential victims will not have their rights violated. In cases of moral blackmail, however, the would-be victims will have their rights violated. In the two cases under consideration, there is a second important difference. In case 4 the burden of ensuring the safety of the hostages is borne solely by the owners of the movie only temporarily. When they are compensated for their loss, the burden is distributed evenly. In the case of saving the victims of famine, however, the burden will be borne solely by those whose profits are seized. Such an action would be unfair, and if we permitted it one's property rights would become useless.

Let us consider another pair of cases to illustrate further the difference between situations of moral blackmail and those where the source of evil is nature. Case 8 provides us with an example where the source of evil is a natural disaster. Some people agree that the observer is permitted to divert the course of the rock in order to save the twenty. Consider an analogous case, though, where the source of evil is a moral blackmailer. Suppose that the blackmailer demands that we kill a certain person or he will kill his twenty hostages. Surely it can be argued that it would be wrong to comply with this demand; note the earlier discussion of the extended version of case 4 and of case 5 (in Section II). Again, we need to ask what justification there is for judging these cases differently. The key to the difference between case 8 and the analogous case involving moral blackmail is an understanding of the rights of the would-be victims. As case 8 is described, neither the twenty nor the two have more of a claim against being crushed by the rock. Because of this, the observer may act to save the greater number. One might wonder, then, whether all cases involving natural evils should be handled by utilitarian reasoning. It can be shown, I think, that they should not. We can imagine a more detailed description of case 8 that would suggest that one group had a stronger claim than the other. Suppose that the twenty were in an area marked *Danger—Falling Rocks* and that the two were in a safety zone. In such a situation, I think, the observer is not permitted to divert the course of the rock. But unless such special conditions obtain,

the observer may act to save the twenty.²⁸ In the case of moral blackmail, however, an agent would be violating an innocent person's rights if he killed him. His failure to kill the one will lead to twenty innocent people's being killed, but the source of the threat to them is quite different from the threat he would pose to the one. The source is someone else's evil action. The one person does have a claim against the agent, and his claim is stronger than the claim the twenty have against that same agent to be saved; or rather, his claim is stronger when the agent's saving the twenty will irreparably violate his rights. So there are important differences between cases of moral blackmail and situations involving natural evils.

It might be thought that the crucial feature about cases of moral blackmail is that they involve the immoral actions of others. There are, however, reasons to doubt this. Consider the following case, originally presented by Judith Jarvis Thomson.

CASE 9: Harry is President and has just been told that the Russians have launched an atom bomb toward New York. The only way in which the bomb can be prevented from reaching New York is by deflecting it, but the only deflection path available will take the bomb to Worcester. Harry can do nothing, letting all of New York die; or he can press a button, deflecting the bomb and killing all of Worcester.²⁹

As Thomson remarks, many people think that Harry is permitted to deflect the bomb in this case. This is certainly a situation where the immoral actions of others are involved, but the agents in question are not blackmailing anyone; they are doing their own dirty work. But consider an analogue of case 9 that does involve moral blackmail. If someone demands that we kill the citizens of Worcester or he will kill the citizens of New York, it seems plausible to say that it is wrong for us to comply with this demand even though the act would prevent an evil exactly as great as the one that Harry can prevent in case 9. I, at least, have different intuitions about these two cases. How might one justify such different judgments about these cases?³⁰ Though in each case the immoral actions of others are involved, the conditions or circumstances of the two cases are different. For reasons that are not always easy to explain, we are especially repulsed by the idea of one moral agent manipulating the other. This, of course, is what happens in cases of moral blackmail. The blackmailer attempts to get a person to do certain acts by threatening to do something much worse. To the extent that the person complies with these demands, he is surrendering his moral

28. My remarks about case 8 follow the same line developed by Judith Jarvis Thomson in "Killing, Letting Die, and the Trolley Problem," *Monist* 59 (1976): 204–17.

29. *Ibid.*, p. 208.

30. Thomson's explanation of the moral difference between these two cases is (presumably) that in situations like case 9 we are doing something to the object of threat, not to people; but in the analogue of that case we would be doing something to people (*ibid.*, p. 216).

autonomy. He is, in a sense, a puppet in the blackmailer's hands. Blackmail situations are characterized by manipulation via threat, and this feature is not present in a situation like case 9. In that case one encounters the actual wrong actions of others as a condition of the world to be taken into account; in cases of moral blackmail the possible wrong acts of others are directed toward an agent as a threat.³¹ Though much more needs to be said about this, this difference does seem to be a morally relevant one. But if the manipulation of the blackmailer is so bad, one might ask, why is it ever permissible to comply with his demands? As long as amends can be made for the (normally wrong) act that the agent performs, giving in to the demands does not seem so bad; or at least it is worth it to prevent a greater evil. If one can restore moral order, the effects of the manipulation are minimal. However, if no amends can be made or if the moral order cannot be restored, the line must be drawn. That degree of moral manipulation cannot be tolerated no matter how great the evil is that the blackmailer will bring about.

If what I have argued is correct, situations of moral blackmail cannot be indiscriminately treated as one among many kinds of cases where the agent can prevent a greater evil only by doing evil. Cases of moral blackmail differ in important ways from those where the source of evil is nature, and surprisingly they also differ from situations involving the immoral actions of other persons not engaged in moral blackmail. As a result, the moral principle(s) governing the actions of the victim of moral blackmail will be different from those governing other situations where one can prevent a greater evil.

VI

Principle (*P*), as it is now stated, indicates the conditions under which it is permissible for an agent to do something evil in order to prevent a moral blackmailer from doing a greater evil. One may wonder, though, whether this version of the principle is not too weak. Should we not say that when the conditions stated in the principle obtain, one is morally required to prevent the greater evil? Consider first the case in which other innocent people suffer when the agent does evil in order to prevent someone else from doing a greater evil. This sort of situation is depicted in case 4. If the agent complies with the demands of the Hanafi Muslims, other innocent persons will be harmed, though not irreparably. In this case it does seem that the agent is obligated, not merely permitted, to do that which is necessary to save the hostages. The rights of the hostages that are at stake are so strong that we ought to do all that

31. In case 9 the immoral action of another is a condition for one's own act in that the action (or perhaps we should say its effects) is already in progress. Another type of situation different from cases of moral blackmail is that in which one knows (or has good reason to believe) that someone else will do something wrong if he gets in a certain position, and one can prevent the act from occurring only by getting in the position oneself and doing something less evil. An excellent example of such a case is presented by Williams, pp. 97–98.

we can to protect them, short of violating another person's equally strong right. But there is also an obligation to make amends to the innocent parties who are harmed. Even the libertarian, who denies that there is a general duty of beneficence, acknowledges that there is an obligation (at least for those acting on behalf of the state) to prevent violations of rights. Protecting one's rights is important enough to allow some evil to be done, provided that those who are harmed are compensated. It does seem, however, that the agent in question should not be required to bear the full burden of compensation. On whom does this obligation fall? Clearly, if possible, the moral blackmailer should be forced to compensate the innocent victim. But if that is not possible, then this is an obligation that should fall on society as a whole. Some particular individual is required to do evil in order to prevent the blackmailer from doing a greater evil because he is in a unique position to do so. This is not normally the case, though, when it comes to compensating those harmed. And there is an obvious sense in which everyone benefits from the practice of protecting one's rights even if something evil must be done. Such a practice will serve to protect better the strongest and most important rights of everyone. So it is plausible to say that, in cases like this, one is required to comply with the demands of the blackmailer.

Consider now the case where the agent who does the evil in order to prevent a greater evil is also the innocent person who will suffer. Case 3 provides such an example. As I describe the case, Isabella must do something which will bring harm to herself so that Claudio's life will be spared. Do we want to say that she is morally required to do the act, or merely that doing such an act is permissible? There seem to be good reasons for one to balk at saying that the agent is obligated to perform such an act. We do have the concept of a supererogatory act, an act that is above and beyond the call of duty. Typically a supererogatory act is one that requires a great sacrifice on the part of the agent, a sacrifice so great that while we characterize the act as one that is morally ideal, we think that it is unreasonable to say that the act is required. Some might claim that in any situation where the agent of an act will be harmed it is implausible to say that the act is morally required. To be sure, it is argued, if the act will prevent a greater evil, it would be a good thing for the agent to do it. But agents cannot be obligated to do acts that will harm themselves; such acts are permissible, but not required. This position, I think, is incorrect. Surely agents are sometimes required to do acts that involve some harm to themselves. In fact, it has been argued that it is implausible to distinguish acts that are obligatory from those that are supererogatory in terms of the degree of harm that is done to the agent because sometimes the performance of duty requires a great sacrifice and some acts of supererogation are trifling.³² The story of Captain Oates provides an example of the former.

32. This is argued by Roderick M. Chisholm in "Supererogation and Offence," *Ratio* 5 (1963): 1-14, see esp. pp. 3-4.

CASE 10: Captain Oates was a member of Robert Scott's expedition. As Oates and his companions were returning from the South Pole, the Captain had an accident and was disabled. The group was far from the depot which was their only source of food and shelter. It was obvious to all that if the others continued to help Oates, none would make it back safely. However, if Oates would detach himself from the group (an act that would ensure a quick death for him), the others would have a chance to return safely. What should Captain Oates do?

As those familiar with this story know, Oates did detach himself from the group, but, tragically, his companions did not survive. It is nevertheless interesting to consider the case as it is stated because it is likely that the evidence available to Oates indicated that his companions might survive; moreover, we can easily imagine a case that differs from the actual one in the respect that the companions do survive, and considering such a case may be instructive. In circumstances such as these, I want to argue, it is reasonable to say that Oates is morally required to detach himself from the group, in spite of the fact that such an act will lead to his death. Though such an act might well be described as heroic, it is still plausible to say that it is his duty and not a supererogatory act. If this is correct, then in some circumstances one can be morally required to make great sacrifices, and so one cannot object to the strengthened version of principle (*P*) on the general grounds that no agent can be required to do something that will result in harm to himself.

Of course, not everyone agrees with my judgment about the Oates case.³³ The critic might argue that the harm that Oates will suffer is irreparable harm—the end of his life—and any principle which requires an agent to do something which will result in irreparable harm to himself is much too austere. In the abstract, this might seem correct. In this particular case, however, I think that it can be argued that Oates is required to separate himself from the others. It is a crucial and morally relevant fact, I believe, that the same irreparable harm will come to Oates no matter what he does. Whether he leaves the group or not, he will die soon. So just as principle (*P*) forbids one to do evil in order to prevent a greater evil if doing so irreparably violates some other nonconsenting, innocent person's rights, so too it does not normally require an agent to do an act that will result in irreparable harm to himself. But in each case there is an exception. If the party in question—either some innocent person or the agent himself—will suffer the same irreparable harm no matter what the agent does, then this evil must be done to prevent a greater evil.³⁴

33. See, for example, W. D. Hudson, *Modern Moral Philosophy* (Garden City, N.Y.: Doubleday & Co., 1970), pp. 220–22.

34. Two points of explanation must be made. First, one might think that since the situations of Joe (in the modified version of case 7) and Captain Oates are apparently analogous, the judgments I have made in these two cases are inconsistent. There are, however, important differences. In case 7 I was discussing the obligation of others' responding to the demand to kill Joe; in case 10 I am discussing the obligation of Oates. In

In concluding, there are two related objections to the strengthened version of principle (*P*) that must be considered. First, it might be thought that the principle requires too much. Consider case 7, the situation in which the villagers are ordered to kill Joe or the entire village will be destroyed. Suppose, for the sake of simplicity, that the population of the village is twenty adults. Now, the objection goes, surely if all twenty agree to fight the brigands in spite of the odds, they are permitted to do so. But the strengthened version of (*P*) does not allow this, since the same rights of the same person (Joe) will be irreparably violated even if they do not comply with the demands. This objection can be answered, however. The new principle can be stated so that one is required to prevent the blackmailer from doing a greater evil only if his original victims (i.e., those whom he threatens to harm if one does not comply) are nonconsenting. If his would-be victims agree to fight it out to the bitter end, they have a right to do so. In this case the villagers are the potential victims, and if they prefer to oppose the blackmailer rather than kill Joe, they may do so.

The second objection is that the principle forbids too much. Suppose in the extended version of case 4 that the Black Muslims volunteer to be turned over to the Hanafis in order to save the hostages. If we comply with the demand, they will be harmed irreparably. But this would not be the case if we failed to comply with the demand. So the strengthened version of principle (*P*) forbids us to comply with the demand. Yet surely if the Black Muslims volunteer to be traded for the 130 hostages, it would be permissible to comply. The response to this objection is that the strengthened version of principle (*P*) does not forbid us to comply with the blackmailer's demands in these circumstances.³⁵ Both clauses of the principle are stated so that one is allowed to prevent a greater evil if the innocent parties (other than the original victims) whose rights will be violated consent to this. The principle does not forbid such acts of herosim.

The question discussed in this essay is to what extent, if any, the actions of a moral blackmailer might affect one's actual obligations. It has been argued that each of the extreme theories—the Kantian theory and the dirty-hands thesis—is inadequate. Each of these views overlooks some morally relevant factors. A theory which does not ignore these factors and which can be defended on independent grounds is a strengthened version of principle (*P*), principle (*P'*), which we may now state as follows:

addition, case 7 is one of moral blackmail and case 10 is not, and, as I have argued, that sometimes makes a moral difference. Second, Isabella ought to comply with the demand of Angelo only on the assumption that the harm she will suffer is not irreparable. I realize that victims of sexual exploitation sometimes do suffer irreparably, and Isabella might be particularly susceptible to this.

35. This act will be permitted, though, only if it really does prevent a greater evil. Again, we will have to take into account the possible bad consequences of setting a precedent of this sort.

An agent in a situation of moral blackmail is morally required to do something evil in order to prevent the blackmailer from doing a greater evil (to nonconsenting persons) just in case either (1) the agent's action does not irreparably violate the rights of some nonconsenting, innocent person, and if an innocent person's rights are violated he is to be compensated, or (2) if a nonconsenting, innocent person's rights are irreparably violated by the agent's action, then the same (or equally strong) rights of that same person would have been violated even if the agent did not comply with the blackmailer's demand.